



Paper Type: Research Paper

A Hybrid Approach to Analyzing Factors Influencing International Student Retention in Turkish Higher Education: Integrating Fuzzy DEMATEL, Machine Learning, and Statistical Methods

Serra Begum Usta Ergun* 

Artificial Intelligence Research and Application Center, Istinye University, Istanbul, Türkiye; begum.usta@istinye.edu.tr.

Citation:

Received: 17 August 2024
Revised: 19 November 2024
Accepted: 09 January 2025

Usta Ergun, S. B. (2025). A hybrid approach to analyzing factors influencing international student retention in Turkish higher education: integrating fuzzy DEMATEL, machine learning, and statistical methods. *Journal of fuzzy extension and applications*, 6(2), 233-259.


Abstract


Retention of international students poses a significant challenge for higher education institutions due to its social and economic consequences. This research investigates the connections between important variables, including intention to leave, integration commitment, educational commitment, student satisfaction, and challenges faced. A comprehensive model was utilized to examine how these factors influence retention, with a specific focus on the impact of demographic factors. Data was gathered from 2,736 international students enrolled in universities in Turkey and analyzed using IBM SPSS 26 and R software. The findings reveal that the intention to leave exhibits a negative correlation with integration commitment, educational commitment, and satisfaction, while showing a positive association with encountered difficulties. Furthermore, demographic factors such as language skills, quality of life, and family support play a significant role in influencing retention. The results of the hypothesis testing were further substantiated through the use of the fuzzy DEMATEL method, reinforcing the connections among the key variables. The outcomes indicate that the choice to leave is a complex, multifactorial issue, significantly influenced by both personal and institutional factors. As a result, it is essential to develop strategies that assist international students in their academic, cultural, and social environments to improve retention.

Keywords: International student retention, Fuzzy DEMATEL, Regression analysis, Student satisfaction.

1 | Introduction

In today's increasingly globalized world, international students play a crucial role in developing higher education systems. Their contributions extend beyond academia, promoting economic and social

 Corresponding Author: begum.usta@istinye.edu.tr

 <https://doi.org/10.22105/jfea.2024.484077.1668>



Licensee System Analytics. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0>).

development in host countries. Investments in higher education, particularly in the recruitment and retention of international students, have been shown to foster significant socio-economic benefits [1], [2].

Over the past two decades, Türkiye has experienced a marked increase in international student enrollment, reflecting a global trend in higher education internationalization. By 2019, the number of international students in Türkiye reached 185,001, signaling a substantial leap compared to previous years. However, along with this growth comes the challenge of retaining these students within Turkish institutions. Research shows that international students often face difficulties that can lead to higher dropout rates [3], [4].

Internationalization has become a cornerstone of higher education policy in many developed countries, where strategies aimed at enhancing student retention have gained significant traction. Studies have highlighted the importance of academic support, social integration, and financial aid in improving student retention [5], [6]. However, despite the increasing numbers of international students in Türkiye, the country has a relatively low student sense of belonging compared to other OECD countries, further underscoring the need for focused retention strategies [7], [8].

This study aims to fill this gap by examining the factors that influence the retention of international students in Turkish higher education. Using a holistic model, it explores the relationships between intention to leave, integration commitment, educational commitment, student satisfaction, and demographic variables. By identifying the key predictors of student retention, this research seeks to provide actionable insights for policymakers and institutions to enhance the retention of international students in Türkiye.

2 | Problem Statement

In the last two decades, the retention of international students has gained prominence as a key indicator of institutional success in higher education. Universities around the world have recognized that the ability to retain students is not only essential for maintaining enrollment numbers but also for fostering long-term academic success, cultural exchange, and economic benefits. Retention is influenced by a wide array of factors, including students' academic performance, social integration, financial support, and institutional engagement. In Türkiye, as the number of international students continues to grow, understanding the dynamics of student retention has become crucial for institutions seeking to improve their global competitiveness and provide a supportive learning environment.

Despite the rapid increase in the number of international students in Türkiye, many face significant obstacles that hinder their ability to integrate into the academic environment fully. Cultural differences often create challenges in building connections with peers and faculty, leading to feelings of isolation. Language barriers further exacerbate these difficulties, limiting students' participation in classes and social activities. Additionally, Turkish higher education institutions frequently lack comprehensive support structures specifically designed for the diverse needs of international students, such as targeted mentorship programs and specialized counseling services. These challenges contribute to higher dropout rates, positioning Türkiye behind more established international education systems. The urgency of addressing these issues is clear, as effective retention strategies are crucial for maintaining the country's competitive edge in the global academic landscape and ensuring a positive experience for international students.

The following section delves into the concept of student retention and the various factors that influence it, offering a comprehensive overview of the theoretical frameworks and empirical evidence that guide retention strategies in higher education.

2.1 | The Concept and Determinants of Student Retention

Student retention refers to the ability of educational institutions to retain students until they complete their academic programs. This concept is particularly important for higher education institutions, where the retention of international students is a key factor in institutional success, as it impacts academic outcomes, financial stability, and global reputation. In the context of international education, retention encompasses a

variety of efforts aimed at ensuring that students remain engaged in their studies and are supported academically, socially, and culturally throughout their educational journey.

Several factors influence student retention, especially for international students who may face additional challenges. These factors can be broadly categorized into academic, social, financial, and institutional domains. Academic factors include students' performance, access to academic support, and alignment between their expectations and the reality of their educational experience. Institutions that provide effective academic counseling, tutoring services, and career guidance tend to have higher retention rates, as students feel more supported in achieving their academic goals.

Social factors also play a critical role in retention. For international students, integration into the social fabric of the institution and the local community is essential. Research has shown that students who are well-integrated socially, whether through campus organizations, peer relationships, or cultural exchange programs, are more likely to persist in their studies. Tinto's theory of social and academic integration [9] suggests that a strong sense of belonging on campus leads to higher retention rates, a finding supported by many studies.

Financial factors are another determinant, particularly for international students who may face additional financial burdens, such as higher tuition fees and living expenses. Access to scholarships, grants, and work opportunities can ease these financial pressures, helping students to focus on their studies without the added stress of financial instability. As highlighted in previous studies [10], financial support plays a significant role in preventing dropouts among economically disadvantaged students.

Institutional factors include the quality of services offered by the university, such as health and wellness support, housing, and administrative assistance. Institutions with a strong support system that addresses both the academic and non-academic needs of students are more likely to retain their international students. Furthermore, an institution's commitment to diversity and inclusivity, ensuring that international students feel welcomed and valued, directly influences their overall satisfaction and retention.

Demographic factors also contribute to student retention. Variables such as age, gender, family income, and prior academic preparation can affect a student's ability to persist in higher education. For international students, language barriers, cultural adjustment, and differences in educational systems add layers of complexity that can affect their retention. The interplay of these demographic and institutional factors creates a multifaceted environment where retention is not solely dependent on one aspect but on the combination of several influences.

In Turkish higher education, where the number of international students has been steadily increasing, institutions must recognize these diverse factors and develop targeted retention strategies. As research suggests [7], [8] enhancing student engagement through academic support, social integration, and financial assistance is key to improving retention rates. By addressing these determinants holistically, institutions can create a supportive environment that promotes long-term academic success for international students.

2.2 | Student Retention Models

Research on student retention has been extensive, with numerous models developed to explain the factors that contribute to student persistence in higher education. The most recognized of these is Tinto's student integration model [9], which emphasizes the significance of both academic and social integration in determining whether a student persists or drops out. According to Tinto's model [9], students' pre-university characteristics, such as family background, prior educational experiences, and personal attributes, shape their initial educational commitments and expectations upon entering higher education. These expectations evolve as students interact with the academic and social environments of their institutions. Tinto [11] later expanded on his theory, stressing that social and academic adaptation should be viewed as interdependent processes rather than separate phenomena.

Another key model is Pascarella and Terenzini's longitudinal model [12], which extends Tinto's [11] framework by examining how student interactions with faculty, peers, and the broader campus environment

influence their academic success and persistence. This model highlights the importance of continuous interaction between students and the institutional environment throughout their educational journey, stressing that these interactions help shape academic outcomes and retention rates.

Bean and Metzner [13] shifted the focus to non-traditional students, arguing that external factors such as family responsibilities, employment, and financial constraints play a larger role in their retention decisions than in traditional students. This model emphasizes that non-traditional students often have lower levels of campus involvement and are more heavily influenced by external pressures, making targeted support services essential for improving their retention rates.

Additionally, Astin's theory of student engagement [14] introduces the concept that student retention is closely linked to the amount of physical and psychological energy that students invest in their studies and institutional activities. According to Astin [14], factors such as motivation, academic skills, and institutional support play critical roles in determining students' success. His model also underscores the importance of creating opportunities for students to engage both academically and socially, suggesting that students who are more involved in their campus environments tend to have higher retention rates.

These models provide a comprehensive understanding of the multiple dimensions that impact student retention, emphasizing the importance of both internal factors (such as academic commitment and social integration) and external factors (such as family support and financial stability). Together, they form the foundation of contemporary retention strategies that aim to support students holistically throughout their academic journey.

2.3 | Sample Applications of Countries

International student retention is a key concern for higher education institutions worldwide, and understanding the successful strategies employed by different countries can provide valuable insights. In this section, we explore the student retention practices from the United States, Canada, and the United Kingdom, which have consistently attracted and retained large numbers of international students.

Before moving to international case studies, it is important to examine the situation in Türkiye. Turkish universities generally offer academic counseling, orientation programs, academic support services, and financial assistance to international students. However, despite these efforts, Türkiye lacks comprehensive long-term retention strategies specifically tailored for international students. This gap highlights an opportunity for Türkiye to learn from best practices employed in other countries.

2.3.1 | United States of America

The United States has long been a leader in attracting international students, and institutions have implemented various retention strategies to support these students:

- I. Orientation and mentoring programs: Penn state university's orientation program helps international students integrate into the campus community, providing academic and social guidance. Mentorship programs also offer individualized support to students during their transition to university life.
- II. Social and academic support: universities like the university of Southern California and the university of Illinois focus on blending academic resources with social support to foster a sense of belonging. For instance, the Illinois Promise program provides social and financial support to underrepresented students, including international students.
- III. Career preparation: institutions like Northeastern university and Georgia state university emphasize career preparation as a retention tool. Programs such as co-op placements help international students gain practical work experience, enhancing their post-graduation prospects and increasing their likelihood of completing their studies.

- IV. Data-driven retention strategies: the National Center for Education Statistics (NCESs) provides institutions with comprehensive data on retention and graduation rates, enabling universities to evaluate their performance and develop targeted retention strategies [15].

2.3.2 | Canada

Canada's inclusive policies and support structures for international students have made it a top destination for higher education.

Integration programs

The university of British Columbia's international student initiative focuses on assisting students with academic and social integration, ensuring they adapt to both university life and Canadian culture.

Targeted academic and language support

The university of Toronto's Green Path program is an example of how targeted academic and language preparation can mitigate the challenges international students face, particularly those from non-English-speaking backgrounds.

Mentorship and cultural awareness

The university of Windsor offers a comprehensive support system, including cultural awareness training and academic counseling. Toronto Metropolitan university's Tri-Mentoring program pairs new students with experienced peers, creating a network of support that fosters both academic and social success.

2.3.3 | United Kingdom

A variety of successful strategies and programs have been implemented to increase student retention rates in higher education in the United Kingdom. The university of Hull's¹ Go Connect program focuses on improving the social integration and employability skills of international students. The university of Leeds' Plus program offers mentoring, career guidance, and financial support to students from low socioeconomic backgrounds. The university of Birmingham's transition review program, highlighted in Schulmann's study, facilitates the adjustment of first-year students to university life. The university of Edinburgh's student support and personal tutoring program, recognized as a best practice by BeMo in 2024, provides a comprehensive support system with a personal counselor assigned to each student. Meanwhile, the university of Bristol's student inclusion and Wellbeing program aims to enhance students' overall well-being and integration into campus life, offering a range of services such as mental health support, financial assistance, and academic counseling. These programs in the United Kingdom are tailored to meet student needs and are designed to improve graduation rates by enriching the overall university experience.

To better understand the complex relationships between the factors influencing international student retention, this study incorporates Fuzzy Set theory, which allows for handling uncertainty and imprecision in data analysis. Fuzzy logic is particularly useful when dealing with variables that are not strictly binary but instead have degrees of membership, making it suitable for modeling real-world situations where clear boundaries are often lacking. The following subsections provide an overview of key fuzzy concepts, including Triangular Fuzzy Numbers (TFNs), fuzzy arithmetic, and defuzzification, all of which are essential for applying the fuzzy DEMATEL method used in this research.

2.4 | Introduction to Fuzzy Concepts

Fuzzy Set theory, introduced by Zadeh [16], is a mathematical framework designed to handle the concept of partial truth, where elements can belong to a set with varying degrees of membership, rather than the binary membership of classical set theory (i.e., either completely in or out of a set). In many real-world situations,

¹ <https://www.hull.ac.uk/>

boundaries between categories are not sharply defined. For instance, terms like tall or expensive are subjective and vary from one context to another. Fuzzy Set theory provides a structured way to represent and reason with such ambiguity.

In fuzzy sets, the degree of membership of an element is expressed by a value between 0 and 1. A membership value of 0 means the element does not belong to the set, a value of 1 indicates full membership, and any value between 0 and 1 represents partial membership. This flexibility makes fuzzy sets especially useful in modeling uncertainty, vagueness, and imprecise data, where traditional binary logic fails to capture the complexity of the situation.

2.4.1 | Triangular fuzzy numbers

A TFN $\tilde{A} = (a_1, a_2, a_3)$ is defined by a triplet of real numbers $a_1 < a_2 < a_3$ where: 1) a_1 : the lower bound, representing the smallest possible value, 2) a_2 : the peak, representing the most likely value, and 3) a_3 : the upper bound, representing the largest possible value.

The membership function $\mu_{\tilde{A}}(x)$ of the TFN \tilde{A} is given by:

$$\mu_{\tilde{A}}(x) = \begin{cases} \frac{x - a_1}{a_2 - a_1}, & a_1 \leq x < a_2, \\ 1, & x = a_2, \\ \frac{a_3 - x}{a_3 - a_2}, & a_2 < x \leq a_3, \\ 0, & \text{otherwies.} \end{cases}$$

This membership function forms a triangular shape, hence the name TFN.

2.4.2 | Arithmetic operations on fuzzy numbers

Fuzzy arithmetic allows for the manipulation of fuzzy numbers. For two TFNs $\tilde{A} = (a_1, a_2, a_3)$ and $\tilde{B} = (b_1, b_2, b_3)$, the basic arithmetic operations are defined as follows:

I. Addition

$$\tilde{A} + \tilde{B} = (a_1 + b_1, a_2 + b_2, a_3 + b_3),$$

II. Subtraction

$$\tilde{A} - \tilde{B} = (a_1 - b_3, a_2 - b_2, a_3 + b_1),$$

III. Multiplication (by a scalar λ)

$$\lambda \cdot \tilde{A} = (\lambda a_1, \lambda a_2, \lambda a_3), \lambda > 0,$$

$$\lambda \cdot \tilde{A} = (\lambda a_3, \lambda a_2, \lambda a_1), \lambda < 0.$$

These operations maintain the triangular shape of the fuzzy numbers, allowing for consistency in computations involving uncertainty [17-19].

2.4.3 | Defuzzification

When working with fuzzy numbers, it is often necessary to convert these fuzzy values into a single, crisp output for practical decision-making. This process, known as defuzzification, translates the ambiguity of fuzzy numbers into a clear, actionable result. Various methods exist for defuzzification, with the modified Center of Gravity (COG) method being one of the most used. In this subsection, the defuzzification process is

outlined, with particular emphasis on the COG method, which plays a crucial role in interpreting the results of the fuzzy analysis in this study:

$$A_{\text{COG}} = \frac{a_1 + 4a_2 + a_3}{6}.$$

This method provides a more balanced way to summarize a fuzzy number, giving greater weight to the most likely value (a_2) while still considering the lower and upper bounds (a_1 and a_3) [20].

3 | Methodology

The methodological approach of this study is designed to systematically explore the factors influencing the retention of international students in higher education, particularly within the context of Turkish universities. Understanding the factors contributing to the intention to leave or remain enrolled requires a combination of quantitative analysis and a structured approach to data collection. The methodology employed in this study incorporates both descriptive and inferential statistical methods, as well as machine learning techniques and fuzzy DEMATEL, to comprehensively analyze the data.

The study also draws on existing retention models, such as Tinto's theory of integration [9] and Astin's student engagement theory [14], which emphasize the importance of social and academic factors in influencing student retention. By leveraging these theoretical frameworks, this study aims to test hypotheses related to international students' individual difficulties, commitments, and satisfaction levels and how these factors impact their intention to leave. In addition, demographic variables such as family income, gender, and accommodation type are included to provide deeper insights into how these factors may contribute to student retention.

The following subsections outline the purpose of the study, hypotheses, and research model, followed by the data collection methods and analytical techniques employed.

3.1 | Purpose of the Study

The primary objective of this study is to investigate the factors influencing the retention of international students in Turkish higher education institutions. Specifically, the study examines how variables such as individual difficulties, integration commitment, educational commitment, and perceived satisfaction affect students' intention to leave. Additionally, the study explores the impact of demographic factors such as gender, faculty, nationality, family income, pocket money, family residence in Istanbul, accommodation type, and educational status on students' intention to leave their studies.

Through this analysis, the study aims to provide a comprehensive understanding of the challenges faced by international students and to offer insights into how universities can improve their retention strategies. The research is designed to fill the gap in the current literature on international student retention in Türkiye by offering both theoretical and practical implications.

3.2 | Hypotheses of the Research

Nine hypotheses are tested in the research:

- I. H₁: individual difficulties and obstacles positively affect the intention to leave.
- II. H₂: individual difficulties and obstacles negatively affect integration commitment.
- III. H₃: individual difficulties and obstacles negatively affect educational commitment.
- IV. H₄: individual difficulties and obstacles negatively affect perceived satisfaction.
- V. H₅: integration commitment negatively affects the intention to leave.
- VI. H₆: educational commitment negatively affects the intention to leave.

- VII. H₇: perceived satisfaction negatively affects the intention to leave.
- VIII. H₈: there is a positive relationship between integration commitment, educational commitment, and perceived satisfaction.
- IX. H₉: intention to leave varies according to demographic variables such as gender, family income, accommodation type, nationality, and other demographic factors.

3.3 | Research Model

The research model is presented in *Fig. 1*. *Fig. 1* illustrates the conceptual framework of the study, highlighting the interconnections among key factors influencing international student retention. *Fig. 1* shows how individual difficulties/barriers impact integration commitment, educational commitment, and perceived satisfaction (H₂, H₃, H₄), which in turn influence the intention to leave (H₅, H₆, H₇). Additionally, demographic variables such as gender, nationality, and family income are shown to play a moderating role in shaping the intention to leave (H₉). The arrows indicate the hypothesized directional relationships, emphasizing the complex interplay between these variables. This visual representation is essential in understanding the comprehensive nature of retention dynamics, allowing for the identification of significant paths and the prioritization of strategic interventions for enhancing student retention.

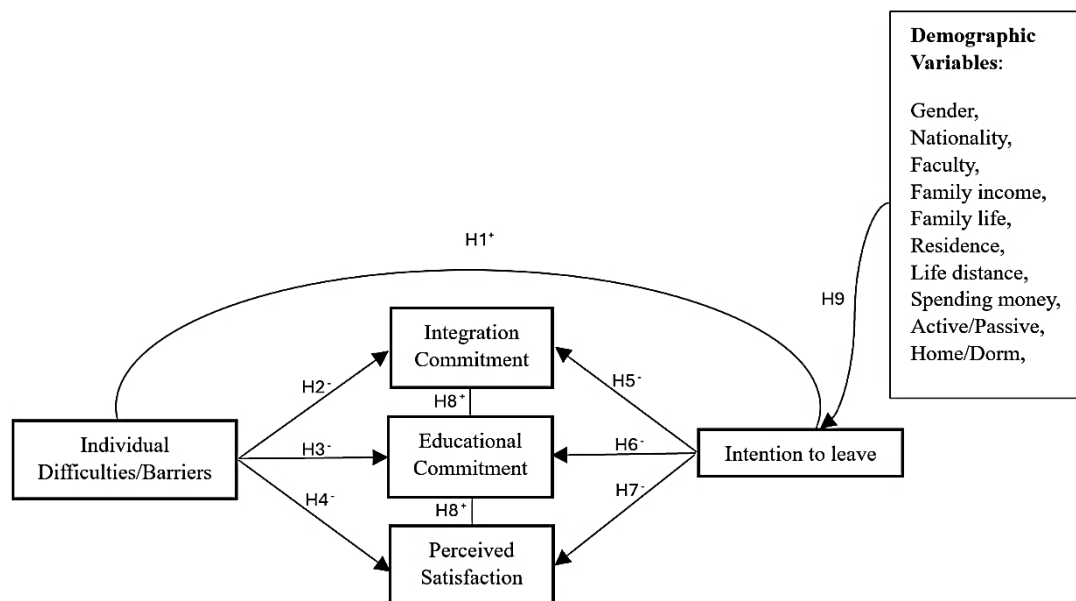


Fig. 1. Research model.

The arrows and directional cues highlight the influence paths, aiding in visualizing how each factor interacts within the retention model.

Individual difficulties/barriers

These refer to challenges faced by students, such as financial problems, language difficulties, or personal issues. According to the model, individual difficulties have a direct influence on several other variables, including intention to leave, integration commitment, educational commitment, and perceived satisfaction. The hypothesis suggests that as difficulties increase, the intention to leave will rise while the other variables will decrease.

Integration commitment

This variable represents the level to which students feel integrated socially and academically within the institution. According to the model, it negatively affects the intention to leave, implying that stronger integration decreases the likelihood of a student leaving. It is also positively correlated with perceived

satisfaction and educational commitment, meaning well-integrated students are more satisfied and committed to their education.

Educational commitment

This measures a student's dedication to their studies and long-term academic goals. The model shows that higher educational commitment is linked to lower intention to leave and greater perceived satisfaction. Educational commitment also reflects students' expectations and experiences in relation to the academic environment.

Perceived satisfaction

Perceived satisfaction reflects how content students are with their academic, social, and institutional experiences. It has a strong negative relationship with intention to leave and positive relationships with both integration commitment and educational commitment. Higher satisfaction levels are indicative of a reduced desire to leave the institution.

Demographic variables

these include gender, nationality, faculty, family income, family life, residence, life distance, spending money, active/passive, home/dorm, and other demographic factors. the model suggests that these variables have a direct influence on the intention to leave but do not necessarily affect the other variables in the same way.

Relationships between variables

The central variable of interest is intention to leave, which is influenced by multiple factors:

- I. Positive effect from individual difficulties: as challenges grow, students are more likely to consider leaving.
- II. Negative effects from integration commitment, educational commitment, and perceived satisfaction: the stronger the students' commitment and satisfaction, the less likely they are to leave.
- III. Interconnections: the model also illustrates the interrelatedness between integration commitment, educational commitment, and perceived satisfaction. students who are more integrated into the academic and social environment of their institution are likely to be more satisfied and committed to their studies.

The relationships between the identified factors were established based on a comprehensive approach combining theoretical frameworks and empirical data analysis. Initially, a thorough literature review was conducted to identify potential connections among variables, such as integration commitment, educational commitment, and perceived satisfaction. These connections were further refined through structured expert consultations, where higher education specialists provided their evaluations. To validate these relationships, statistical analyses such as Pearson correlation and multiple regression analysis were performed, assessing the strength and direction of interactions between variables. Additionally, the fuzzy DEMATEL method was employed to map and quantify the causal relationships, revealing both direct and indirect impacts among factors. This multi-layered approach ensured that the defined relationships were backed by both qualitative insights and quantitative validation, enhancing the methodological rigor of the study.

3.4 | Sample and Data Collection Method

The sample for this study consists of 2,736 international students enrolled in a foundation university. These students represent both active and passive statuses, as well as those who have canceled their registrations. Data collection was conducted through two methods: 1) the university's student information system for active and passive students, and 2) MS forms for those who had canceled their registration as of May 2024.

A total of 2,145 active students, 268 passive students, and 323 students who had canceled their registration participated in the study. This combination provides a comprehensive dataset that captures a wide range of student experiences and statuses. Collecting data from multiple sources ensures that both current and former students are represented, giving the study depth in examining the factors influencing student retention.

The data collected includes demographic information, such as gender, nationality, family income, and residential information. Additionally, it captures academic performance, integration commitment, educational commitment, individual difficulties faced, and perceived satisfaction, all of which are critical variables in understanding the students' intentions to leave the institution.

By utilizing both internal systems and external survey tools, this study achieves a robust and representative sample, which is essential for providing reliable and generalizable findings on the factors affecting international student retention in Turkish higher education.

3.5 | Analysis Method

The data obtained in the study were analyzed using the SPSS and R statistical programs. A significance level of $p < 0.05$ was adopted for all statistical tests. The following steps were taken during the analysis:

3.5.1 | Descriptive statistics

Initially, descriptive statistics were calculated for the demographic variables and key measures, such as intention to leave, individual difficulties, integration commitment, educational commitment, and perceived satisfaction. This provided a summary of the general characteristics of the sample and the distribution of responses across different variables.

3.5.2 | Difference tests

To examine whether the dependent variable, intention to leave, differed according to various demographic factors, an independent sample t-test was applied for two-category variables, such as gender, and a one-way Analysis of Variance (ANOVA) was conducted for variables with more than two categories, such as family income and faculty. These tests helped identify statistically significant differences in the intention to leave based on demographic characteristics.

3.5.3 | Correlation analysis

Pearson correlation analysis was conducted to determine the direction and strength of the relationships between the study's key variables. This analysis provided insights into how strongly variables such as individual difficulties, integration commitment, and perceived satisfaction were related to the intention to leave.

3.5.4 | Regression analysis

Multiple linear regression analysis was employed to assess the impact of independent variables (such as individual difficulties, integration commitment, educational commitment, and perceived satisfaction) on the dependent variable (intention to leave). The results of this analysis helped to quantify the effect of each factor on students' intention to leave.

3.5.5 | Machine learning algorithms

To further explore the factors influencing student retention, machine learning algorithms were applied. These included k-Nearest Neighbor (k-NN), Naive Bayes, Logistic Regression, Support Vector Machines (SVMs), and Random Forest. The data was divided into 80% training and 20% test data to evaluate the models' performance using metrics such as accuracy, precision, recall, and F1-score. The application of machine learning allowed for predictive modeling of student retention, identifying key factors and their relative importance in explaining the students' intention to leave.

3.5.6 | Model performance evaluation

The models were evaluated using the test dataset, and the performance of each machine learning algorithm was compared. Random Forest was found to provide the best performance, with the highest accuracy in predicting student retention outcomes.

3.5.7 | Fuzzy DEMATEL method

The fuzzy DEMATEL methodology was employed to analyze the interrelationships and influence of various criteria on project success under conditions of uncertainty. By incorporating fuzzy logic into the traditional DEMATEL approach, it better handles vagueness and subjectivity in expert judgments. The method helps visualize cause-effect relationships by constructing a fuzzy influence matrix through pairwise comparisons. The analysis revealed the direct and indirect impacts of the criteria, highlighting the most influential and dependent ones in a more nuanced way, accounting for ambiguity in the data [21], [22].

Fuzzy DEMATEL has been widely used in various fields to model complex decision-making processes involving uncertainty. For instance, studies have employed fuzzy DEMATEL to assess educational quality and prioritize factors influencing academic performance (e.g., Gül [23], Torkzadeh et al. [24]). These studies demonstrate how fuzzy DEMATEL and its extensions can be applied to understand complex interdependencies and prioritize influential factors in educational contexts. In non-educational contexts, it has been applied to areas such as sustainable supply chain risk evaluation, showcasing its utility in analyzing multi-criteria decision-making scenarios (e.g., Ansari et al. [25]). These applications underscore its robustness in analyzing interrelationships among variables, making it well-suited for this study on student retention.

To provide readers with a better understanding of the historical background and primary applications of the fuzzy DEMATEL method, *Table 1* outlines its key developments and notable uses over time. This summary emphasizes its evolution and versatility in tackling complex decision-making challenges in both educational and non-educational settings.

Table 1. Key developments and notable uses over time.

Year	Title	Authors	Description	Application Area
2019	A hybrid fuzzy DEMATEL-AHP/VIKOR method for LMS selection	Ayouni et al. [26]	Combined fuzzy DEMATEL with AHP/VIKOR for selecting learning management systems.	Educational technology selection
2020	Defining the strategic impact-relation map for the innovative investments based on IT2 fuzzy DEMATEL	Diñçer and Yüksel [27]	Used IT2 Fuzzy DEMATEL for mapping strategic relations in educational investments.	Strategic planning in education
2020	Adapting to PSTs' pedagogical changes in sustainable mathematics education through flipped E-Learning: ranking its criteria with MCDA/F-DEMATEL	Jeong and González-Gómez [28]	Applied fuzzy DEMATEL to rank criteria for sustainable mathematics education.	Flipped classroom and mathematics education
2022	Picture fuzzy extension of DEMATEL and its usage in educational quality evaluation	Gül [23]	Extended DEMATEL with picture fuzzy sets to assess educational quality.	Educational quality evaluation
2022	To determine the critical factors for the adoption of cloud computing in the educational sector in developing countries – a fuzzy DEMATEL approach	Thavi et al. [29]	Identified and analyzed critical factors for cloud computing adoption using fuzzy DEMATEL.	Cloud computing in education

Table 1. Continued.

Year	Title	Authors	Description	Application Area
2023	Multi-Criteria Decision Analysis and Fuzzy-Decision-Making Trial and Evaluation Laboratory (MCDA and F-DEMATEL) method for flipped and sustainable mathematics teaching	Jeong et al. [30]	Integrated fuzzy DEMATEL with MCDA for evaluating flipped teaching methods.	Mathematics and sustainable education
2023	Analysis of the organizational knowledge management system of nurse education centers with hybrid fuzzy DEMATEL-network DEA method	Banihashemi et al. [31]	Evaluated knowledge management systems in nursing education using a hybrid method.	Nursing education management
2024	A q-Rung Orthopair Fuzzy Set Extension of the DEMATEL and its application in the education sector	Revalde et al. [32]	Introduced q-Rung Orthopair Fuzzy DEMATEL for educational applications.	Advanced fuzzy logic in education
2024	Influencing factors of spatial ability for architecture and interior design students: a fuzzy DEMATEL and interpretive structural model	Amro and Dawoud [33]	Analyzed factors affecting spatial ability in architecture students.	Architecture and design education

The fuzzy DEMATEL method involves several key steps to analyze the interrelationships between factors under conditions of uncertainty. Below is a detailed description of these steps, including the conversion of linguistic terms into TFNs:

Define the problem and identify factors

The first step involves defining the problem and identifying the factors or criteria that influence the system under study. In this case, the focus is on factors influencing international student retention.

Construct the fuzzy direct-relation matrix

A group of experts evaluates the direct influence of each factor on the others using linguistic terms (e.g., low, medium, high). These linguistic terms are converted into TFNs to represent the degree of influence, capturing the inherent uncertainty in expert judgments. Below is the conversion *Table 2*.

Table 2. Linguistic terms to triangular fuzzy numbers.

Linguistic Term	TFNs
Very Low (VL)	(0, 0, 0.25)
Low (L)	(0, 0.25, 0.5)
Medium-Low (ML)	(0.25, 0.5, 0.75)
Medium (M)	(0.5, 0.75, 1)
Medium-High (MH)	(0.75, 1, 1.25)
High (H)	(1, 1.25, 1.5)
Very High (VH)	(1.25, 1.5, 1.75)

In this step, where the fuzzy direct-relation matrix is constructed, it is essential to combine the evaluations of multiple experts into a single consensus value for each pairwise comparison. A common approach is to use the arithmetic mean of the TFNs provided by the experts. In fuzzy logic, rather than having a crisp or exact

value, the influence of one factor on another is represented by a range of possible values. For each triangular fuzzy $\tilde{x}_{ij}^k = (a_{ij}^{l,k}, b_{ij}^{m,k}, c_{ij}^{u,k})$ given by the k -th expert, the aggregated fuzzy number $(\bar{x}_{ij}^l, \bar{x}_{ij}^m, \bar{x}_{ij}^u)$ is a way to capture the uncertainty and variability in the expert's opinion about the influence, and it is calculated as follows:

$$\bar{x}_{ij}^l = \frac{\sum_{k=1}^n a_{ij}^{l,k}}{n}, \bar{x}_{ij}^m = \frac{\sum_{k=1}^n b_{ij}^{m,k}}{n}, \bar{x}_{ij}^u = \frac{\sum_{k=1}^n c_{ij}^{u,k}}{n},$$

where 1) n : the number of experts, 2) k : index of experts, 3) the lower bound $(a_{ij}^{l,k})$ accounts for a conservative estimate, 4) the median value $(b_{ij}^{m,k})$ captures the expert's most likely assessment, and 5) the upper bound $(c_{ij}^{u,k})$ reflects a more liberal or optimistic assessment of influence.

By combining these three values, the fuzzy DEMATEL method allows for a more flexible and comprehensive representation of expert judgments, acknowledging the inherent uncertainty in human assessments.

This TFN will later be defuzzified or combined with other expert inputs to obtain a clearer view of the influence factor, which helps to build the overall direct-relation matrix in the fuzzy DEMATEL process.

Normalize the fuzzy direct-relation matrix

In this step, the fuzzy direct-relation matrix is normalized to ensure that all values lie within a standard range, typically between 0 and 1. The normalization process for each element of the matrix \tilde{d}_{ij} (where represents the fuzzy numbers: lower, middle, and upper bounds) is as follows:

$$\begin{aligned} \tilde{d}_{ij}^l &= \frac{x_{ij}^l - \min\{x_{ij}^l\}}{\max\{x_{ij}^l\} - \min\{x_{ij}^l\}}, \quad i, j = 1, \dots, n, \\ \tilde{d}_{ij}^m &= \frac{x_{ij}^m - \min\{x_{ij}^m\}}{\max\{x_{ij}^m\} - \min\{x_{ij}^m\}}, \quad i, j = 1, \dots, n, \\ \tilde{d}_{ij}^u &= \frac{x_{ij}^u - \min\{x_{ij}^u\}}{\max\{x_{ij}^u\} - \min\{x_{ij}^u\}}, \quad i, j = 1, \dots, n. \end{aligned}$$

The fuzzy matrix \tilde{D} is represented as follows:

$$\tilde{D} = (D_l, D_m, D_u) = ([d_{ij}^l]_{n \times n}, [d_{ij}^m]_{n \times n}, [d_{ij}^u]_{n \times n}).$$

Calculate the fuzzy total-relation matrix

Using the normalized fuzzy direct-relation matrix, the total-relation matrix is calculated. This matrix captures both direct and indirect influences between factors. It is derived using the formula:

$$\tilde{T} = \tilde{D} \otimes (\tilde{I} - \tilde{D})^{-1},$$

where 1) \tilde{T} : fuzzy total-relation matrix, 2) \tilde{D} : normalized fuzzy direct-relation matrix, and 3) \tilde{I} : identity matrix.

The calculation of the inverse matrix $(\tilde{I} - \tilde{D})^{-1}$ follows the methodology outlined by Basaran [34], which is specifically designed for operations involving fuzzy matrices. This method provides a structured approach for

addressing the inherent uncertainty in fuzzy matrices and ensures that the operations are executed with precision.

$$D := \begin{pmatrix} D := D_1^r \sum_{j=1}^n t_{1j}^l \\ D := D_2^r \sum_{j=1}^n t_{2j}^l \\ \vdots \\ D := D_n^r \sum_{j=1}^n t_{nj}^l \end{pmatrix}, \quad r = l, m, u,$$

In each direct-relation matrix T_r , $r = l, m, u$ the sum of the values in a row i of each matrix T_r represents the influence that one specific factor exerts on all other factors. In other words, the sum of the elements in row i (denoted as $D_i^r = \sum_{j=1}^n T_{ij}^r$, $r = l, m, u$ shows the overall influence of factor i).

Besides, in each matrix the sum of the values in a column j of each matrix T_r represents the one specific factor influenced by other factors. In other words, the sum of the elements of column j (denoted as $R_j^r = \sum_{i=1}^n t_{ij}^r$, $r = l, m, u$ shows the overall influenced factor j).

$$R := \begin{pmatrix} R := R_1^r \sum_{j=1}^n t_{1j}^l \\ R := R_2^r \sum_{j=1}^n t_{2j}^l \\ \vdots \\ R := R_n^r \sum_{j=1}^n t_{nj}^l \end{pmatrix}, \quad r = l, m, u.$$

Defuzzification of the matrix

At this stage, the obtained total-relation matrix (T), which contains fuzzy elements, is converted into a matrix with crisp values using a specific method. A commonly used defuzzification technique is the modified COG method

$$TOG = (t_{ij}^l, t_{ij}^m, t_{ij}^u) = \frac{t_{ij}^l + 4t_{ij}^m + t_{ij}^u}{6},$$

where $(t_{ij}^l, t_{ij}^m, t_{ij}^u)$ represents the TFN.

Calculate the $D + R$ and $D - R$ values

For each factor, calculate the $D + R$ and $D - R$ values:

- I. $D + R$ (prominence): represents the total influence a factor has in the system, indicating both how much it influences others and how much it is influenced by them.
- II. $D - R$ (net influence): represents the net influence of a factor. A positive $D - R$ value indicates that the factor acts as an influencer, while a negative $D - R$ value suggests that the factor is more dependent on others.

Determine the threshold and construct the fuzzy influence network diagram

A threshold value is applied to filter out weaker relationships in the total-relation matrix, focusing on the most significant interactions. Based on the threshold, a fuzzy influence network diagram is constructed, visually representing the relationships between factors. Arrows in the diagram indicate the direction and strength of influence between factors.

Interpret the results

The final step involves analyzing the $D + R$ and $D - R$ values, as well as the network diagram, to understand the roles of different factors within the system. This interpretation helps identify which factors are the most critical influencers and which are more dependent, providing valuable insights for decision-making and strategy development.

By integrating both statistical analysis and machine learning techniques, the study provides a comprehensive analysis of the factors influencing the retention of international students in higher education. The findings from these analyses were used to test the study's hypotheses and to develop actionable recommendations for improving student retention strategies.

4 | Finding

This section of the study includes the findings obtained in the analysis of the data collected in the study.

4.1 | Descriptive Statistics

The descriptive statistical analysis results are given in *Table 3*.

Table 3. Descriptive statistics analysis results.

Variable	Attribute	Intention to Leave	Integration Commitment	Academic Commitment	Perceived Satisfaction	Individual Difficulties
Gender	Female	2.799	3.326	3.418	3.151	2.764
	Male	2.813	3.293	3.37	3.168	2.811
Family lives in Istanbul.	Yes	2.763	3.382	3.453	3.262	2.673
	No	2.836	3.262	3.352	3.097	2.853
Family Income	0-50,000	2.812	3.354	3.439	3.194	2.84
	50,001-100,000	2.827	3.224	3.302	3.104	2.78
	100,001-200,000	2.74	3.281	3.34	3.091	2.661
	200,001-500,000	2.897	3.166	3.26	3.085	2.655
	500,001 and above	2.772	3.28	3.375	3.153	2.663
Allowance	0-500	2.789	3.352	3.445	3.181	2.825
	501-1000	2.76	2.792	3.284	3.364	3.155
	1001-1500	2.772	2.926	3.111	3.157	3.026
	1501-2000	2.697	2.862	3.203	3.187	3.036
	2000 and above	2.603	2.875	3.23	3.324	3.167
Residence	Dormitory	3.11	2.965	2.979	2.819	2.972
	House	2.765	3.358	3.45	3.209	2.761
Active/passive status	Active	2.691	3.449	3.558	3.298	2.738
	Passive	2.547	3.55	3.767	3.471	2.789

4.2 | Difference Tests

The t-test results conducted within the scope of the study are given in *Table 4*.

Table 4. T-test results.

Variable	t-Value	p-Value	Significant Difference
Living in Istanbul	-1.992	0.046	Yes
Type of accommodation (house/dormitory)	-6.699	<0.001	Yes
Gender	-0.412	0.680	No
Having a job in Istanbul	-0.519	0.604	No

According to the independent sample t-test results, statistically significant differences were found between the students living in Istanbul ($t = -1.992, p < 0.05$) and accommodation type (home/dormitory) ($t = -6.699, p < 0.001$) and their intention to leave. On the other hand, no significant difference was observed between gender ($t = -0.412, p > 0.05$) and having a job in Istanbul ($t = -0.519, p > 0.05$) and their intention to leave. The ANOVA results conducted within the scope of the study are given in *Table 5*.

Table 5. Results from the analysis of variance.

Variable	F-value	p-Value	Significant Difference
Active/passive status	271.555	<0.001	Yes
Family's annual income	33.518	<0.001	Yes
Faculty	15.032	<0.001	Yes
Age	2.885	<0.001	Yes
Country of residence	1.493	<0.001	Yes
Number of siblings	2.185	0.020	Yes
Distance to university	2.896	0.002	Yes
Class	9.908	<0.001	Yes
Student's allowance	1.507	0.197	No

ANOVA results showed that the student's active/passive status ($F = 271.555, p < 0.001$), family annual income ($F = 33.518, p < 0.001$), faculty ($F = 15.032, p < 0.001$), age ($F = 2.885, p < 0.001$), country of residence ($F = 1.493, p < 0.001$), number of siblings ($F = 2.185, p < 0.05$), distance to university ($F = 2.896, p < 0.01$) and class ($F = 9.908, p < 0.001$) variables created statistically significant differences on the intention to leave. No significant difference was found only between the student's pocket money ($F = 1.507, p > 0.05$) and the intention to leave.

For each demographic variable, the descriptive statistics reveal that male students generally exhibit a slightly higher intention to leave compared to female students, along with marginally lower integration commitment. In terms of educational commitment, perceived satisfaction, and individual difficulties, males show only slight differences when compared to females. Furthermore, students whose families reside in Istanbul tend to demonstrate higher levels of integration commitment, educational commitment, and perceived satisfaction.

They also show a lower average in terms of individual difficulties. As expected, intentions to leave are lower among students in higher income groups (100,001TL-200,000TL and 200,001TL-500,000TL).

Similarly, as the amount of pocket money received by students increases, their intention to leave decreases. For instance, perceived satisfaction is highest among students receiving 2000TL or more in pocket money. Students living at home also show lower intentions to leave and higher levels of integration commitment, educational commitment, and perceived satisfaction compared to those living in dormitories, likely due to greater personal space and comfort. Additionally, active students tend to have a lower intention to leave and exhibit higher integration, educational commitment, and perceived satisfaction than passive students.

4.3 | Relationship Analysis Between Variables

Pearson correlation analysis was conducted to examine the relationships between variables within the scope of the study. The analysis results are given in *Table 6*.

Table 6. Pearson correlation analysis results.

Variables	Individual Difficulties/Barriers	Intention to Leave	Student Integration and Commitment	Educational Commitment	Perceived Satisfaction
Individual difficulties/barriers	1				
Intention to leave	0.335**	1			
Student integration and commitment	-0.266**	-0.581**	1		
Educational commitment	-0.335**	-0.692**	0.772**	1	
Perceived satisfaction	-0.283**	-0.711**	0.745**	0.823**	1

**Points out that $p < 0.01$.

Correlation analysis is a statistical analysis method used to measure the relationship between two or more variables and to evaluate the strength of this relationship. The Pearson correlation coefficient takes a value between -1 and +1. Positive correlation coefficients indicate that there is a directly proportional relationship between variables, while negative correlation coefficients indicate that there is an inversely proportional relationship between variables. A correlation coefficient of zero indicates that there is no relationship between variables, while a coefficient approaching -1 or 1 indicates that the strength of the relationship is increasing. In general, correlation coefficients; 1) very weak relationship between 0.00-0.19, 2) weak relationship between 0.20-0.39, 3) medium relationship between 0.40-0.59, 4) Strong relationship between 0.60-0.79, and 5) very strong relationship between 0.80-1.00 [35].

When *Table 6* is examined, we observed that the relationships between all variables are statistically significant at the 0.01 level ($p < 0.01$). The findings regarding the relationships between the variables are summarized below:

- I. Relationships between the individual difficulties/barrier's variable and other variables: 1) positive moderate relationship with intention to leave ($r = 0.335$), 2) weak negative relationship with student integration and commitment ($r = -0.266$), 3) moderate negative relationship with educational commitment ($r = -0.335$), and 4) weak negative relationship with perceived satisfaction ($r = -0.283$).
- II. Relationships between the intention to leave variable and other variables: 1) negative strong relationship with student integration and commitment ($r = -0.581$), 2) negative strong relationship with educational commitment ($r = -0.692$), and 3) negative strong relationship with perceived satisfaction ($r = -0.711$).
- III. Relationships between student integration and commitment and other variables: 1) positive strong relationship between student integration and commitment and educational commitment ($r = 0.772$), and 2) positive strong relationship between student integration and commitment and perceived satisfaction ($r = 0.745$).
- IV. Relationship between educational commitment and perceived satisfaction variables: very strong positive relationship between educational commitment and perceived satisfaction ($r = 0.823$).

In light of these findings, the variable for individual difficulties and obstacles showed weak to moderate correlations with other variables. In contrast, intention to leave demonstrated strong negative relationships with student integration, educational commitment, and perceived satisfaction. Additionally, strong positive relationships were found between student integration, educational commitment, and perceived satisfaction.

4.4 | Regression Analysis

In the following, multiple linear regression analysis was conducted to examine the effects of individual difficulties and barriers (x_1), integration commitment (x_2), educational commitment (x_3), and perceived

satisfaction (x_4) (independent variables) on intention to leave (y) (dependent variable). Toward this end, we use the following regression model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4,$$

in which β_i are regression coefficient and β_0 is the intercept.

The regression analysis results are presented below:

Table 7. Regression analysis results.

Regression Coefficients (β_i)	
Intercept	$\beta_0 = 4.467$
Individual difficulties and barriers	$\beta_1 = 0.182$
Integration commitment	$\beta_2 = 0.005$
Educational commitment	$\beta_3 = -0.262$
Perceived satisfaction	$\beta_4 = -0.409$
R square	0.552

According to the ANOVA results:

- I. Individual difficulties and barriers have a positive impact on intention to leave ($\beta_1 = 0.182$), indicating that as the individual difficulties and barriers faced by students increase, their intention to leave also rises.
- II. Integration commitment shows no statistically significant effect on intention to leave ($\beta = 0.005$). The negligible coefficient suggests that integration commitment does not meaningfully influence a student's intention to leave.
- III. Educational commitment has a negative effect on the intention to leave ($\beta = -0.262$). This result suggests that as students' educational commitment increases, their intention to leave decreases.
- IV. Perceived satisfaction is the variable with the strongest negative effect on the intention to leave ($\beta = -0.409$). As students' perceived satisfaction increases, their intention to leave decreases significantly.

Overall, the model explains 55.2% of the variance in the dependent variable ($R^2 = 0.552$), indicating that it has moderate explanatory power in predicting intention to leave.

4.5 | Machine Learning Analyses for Retention Variable

In this section, machine learning methods were employed to assess the predictive power of the variables explaining the dependent variable, student retention. The analyses were conducted in two phases.

In the first phase, the goal was to predict students' likelihood of leaving using artificial intelligence algorithms. Data from the independent variables were analyzed to determine which algorithm performed best and to rank the variables according to their importance in predicting student retention.

In the second phase, student demographic data were included in the analysis. This was done to evaluate the significance of demographic factors, which are typically known when the student enrolls, and to understand their role in predicting retention outcomes.

Analysis process:

- I. Data cleaning: removing the attributes titled ID, record type_code_text, date and source that will not be used in the analyses from the dataset.
- II. Separating the data into two parts: separating 80% of the dataset for training and 20% for testing.
- III. Normalizing the data.
- IV. Training the model by applying machine learning algorithms.
- V. Algorithms used: kNN, Bayes, Logistic Regression, SVM, Random Forest. Calculation of model performance with test data.

4.5.1 | Analysis results for features

The results of the first stage of the machine learning analysis are presented in *Table 8*. Upon evaluating the performance of five different algorithms, the Random Forest algorithm was found to deliver the best results, achieving the highest accuracy. The ranking of feature importance derived from this algorithm is displayed in *Table 9*.

Table 8. Analysis results for features.

Algorithm	Class	Precision	Recall	F1-score	Accuracy
KNN	0	0.98	0.93	0.95	0.92
	1	0.62	0.87	0.72	
Bayes	0	0.98	0.88	0.93	0.88
	1	0.51	0.87	0.64	
Logistic regression	0	0.98	0.91	0.95	0.913
	1	0.59	0.87	0.7	
SVM	0	0.98	0.91	0.94	0.9
	1	0.56	0.87	0.68	
Random forest	0	0.96	0.99	0.98	0.96
	1	0.91	0.72	0.8	

The analysis of the performance of five machine learning algorithms, as presented in *Table 8*, indicated that the Random Forest algorithm achieved the highest accuracy at 0.960. This superior performance can be attributed to the algorithm's ensemble nature, which aggregates the results of multiple decision trees to reduce overfitting and improve generalizability. The Random Forest algorithm is also highly effective at handling non-linear relationships and capturing feature importance, as shown in *Table 9*. This robustness makes it an excellent choice for analyzing datasets with complex interdependencies, such as those involving various factors influencing student retention. The findings imply that future studies focusing on predictive modeling in educational research could benefit from employing ensemble methods like Random Forest to achieve reliable and accurate results.

Table 9. Order of importance of features.

Feature	Importance Value
Happiness at the university	0.100189
Recommendation of the university to friends in their home country	0.099719
Academic commitment	0.093997
Likelihood of choosing the same university again if given the chance	0.08256
Perceived satisfaction	0.07236
Determination to continue education	0.039747
University's support in achieving academic goals	0.03774
Satisfaction with the quality of education provided by the university	0.034627
Ability to establish positive relationships with academic advisors/lecturers	0.030089
Willingness to pursue a master's degree at the university if possible	0.027706
Integration commitment	0.025012
Belief in the resolution of problems encountered at the university	0.021893
Ability to pay tuition fees on time	0.01787
Feeling at home at the university	0.017092
Not experiencing language issues in courses	0.014352
Experience at the university meeting expectations	0.014051
Active participation in classes	0.012609
Satisfaction with the adequacy of student support services at the university	0.00827
Considering living in Turkey after graduation	0.008235
Regular participation in social activities at the university	0.008159

4.5.2 | Analyses conducted for demographic data

The analysis results performed for demographic data in the second stage of the machine learning analysis are presented in *Table 10*.

Table 10. Machine learning analysis results conducted for demographic data.

Algorithm	Class	Precision	Recall	F1-score	Accuracy
KNN	0	0.94	0.79	0.86	0.77
	1	0.29	0.63	0.4	
Bayes	0	0.96	0.72	0.82	0.73
	1	0.28	0.78	0.41	
Logistic regression	0	0.94	0.71	0.81	0.71
	1	0.25	0.67	0.36	
SVM	0	0.95	0.72	0.82	0.72
	1	0.26	0.7	0.38	
Random forest	0	0.95	0.99	0.97	0.94
	1	0.88	0.63	0.73	

This dataset includes information available before the student enters university, such as demographic characteristics, which are ranked in order of importance regardless of the institution. This analysis is crucial as it highlights the student's risk of leaving based on these factors.

When the performance of five different machine learning algorithms was evaluated, the Random Forest algorithm produced the best results. The ranking of the most important demographic characteristics obtained through this model is shown in *Table 11*. The analysis revealed that family income, distance from the student's residence to the university, and issues related to obtaining a visa or residence permit were the top three factors influencing a student's intention to leave.

Table 11. Demographic data importance order.

Feature	Importance Value
Annual income of the family	0.215799
Distance from the residence to the university	0.109499
Experiencing issues with obtaining/extending visa/residency permit	0.083291
Amount of allowance received from the family	0.072882
Level of English proficiency	0.062649
Level of Turkish proficiency	0.061075
Living in a dormitory or house	0.030856
Family living in Istanbul	0.02384
Family having a job in Istanbul	0.019273

In light of these findings, the status of whether the hypotheses were accepted or not is summarized in *Table 12*. Accordingly, it was determined that all hypotheses within the scope of the relationships between variables were accepted and that the intention to leave was partially accepted because it did not differ according to all demographic variables.

Table 12. Acceptance status of hypotheses.

Hypotheses	Acceptance Status
H ₁ : individual difficulties and barriers positively affect the intention to leave.	Accepted
H ₂ : individual difficulties and barriers negatively affect integration commitment.	Accepted
H ₃ : individual difficulties and barriers negatively affect academic commitment.	Accepted
H ₄ : individual difficulties and barriers negatively affect perceived satisfaction.	Accepted
H ₅ : integration commitment negatively affects the intention to leave.	Accepted
H ₆ : academic commitment negatively affects the intention to leave.	Accepted
H ₇ : perceived satisfaction negatively affects the intention to leave.	Accepted
H ₈ : there is a positive relationship between integration commitment, academic commitment, and perceived satisfaction.	Accepted
H ₉ : the intention to leave varies according to demographic variables.	Partially accepted

4.6 | Fuzzy DEMATEL

In this subsection, the fuzzy DEMATEL method is applied to analyze the complex interrelationships among factors impacting international student retention under conditions of uncertainty. By utilizing fuzzy DEMATEL, this study captures the uncertainty and subjectivity inherent in expert assessments of the factors

influencing student retention. This allows for a more nuanced analysis, which is critical for addressing the complexities of international student dynamics. The fuzzy DEMATEL approach helps assess both the influence and dependence among factors by applying fuzzy membership functions, transforming subjective judgments into quantifiable data and providing clearer insights into the system's behavior.

The influence and dependence of the factors affecting international student retention are analyzed using the fuzzy DEMATEL method. The five key factors under investigation are intention to leave, individual difficulties and barriers, integration commitment, educational commitment, and perceived satisfaction. By incorporating fuzzy logic, the method addresses uncertainty and subjectivity in expert evaluations, providing a more nuanced analysis of the relationships between these factors.

The direct (D) and indirect (R) influences of these factors were calculated based on fuzzy pairwise comparisons. The results are summarized in the table below, where fuzzy values reflect the degrees of influence and dependence, allowing for a more flexible interpretation of the interrelationships. To better visualize the relationships, a fuzzy threshold was applied to the final fuzzy influence matrix, and the resulting fuzzy network diagram is presented in *Fig. 2*.

Table 13. Fuzzy DEMATEL influence and dependence matrix.

Factor	D (Fuzzy Influence)	R (Fuzzy Dependence)	Defuzzified (R)	Defuzzified (D)	D + R	D – R
Intention to leave	(1.519, 1.575, 1.899)	(0.782, 0.982, 1.182)	1.619	0.982	2.601	0.637
Individual difficulties/barriers	(0.832, 0.962, 1.212)	(1.509, 1.577, 1.899)	0.982	1.619	2.601	-0.637
Integration commitment	(0.633, 0.734, 0.893)	(0.127, 0.246, 0.377)	0.743	0.247	0.991	0.495
Educational commitment	(0.385, 0.473, 0.695)	(0.891, 0.976, 1.151)	0.495	0.991	1.486	-0.495
Perceived satisfaction	(0.633, 0.731, 0.903)	(0.643, 0.731, 0.893)	0.743	0.743	1.486	0

In the context of *Table 13*, D + R values reflect the prominence of each factor, indicating their significance in influencing and being influenced by others. For example, intention to leave and individual difficulties/barriers exhibit the highest D + R values of 2.601, underscoring their critical roles in student retention, even under uncertain conditions.

On the other hand, the D – R values help understand the fuzzy net influence of each factor. Positive fuzzy D – R values, such as those for intention to leave (+0.637) and integration commitment (+0.495), suggest that these factors act mainly as influencers in the system, even when uncertainty is a factor. Conversely, factors like individual difficulties/barriers (-0.637) and educational commitment (-0.495) show negative fuzzy D – R values, indicating that these factors are more dependent on other factors, meaning they are influenced by changes in the system, particularly when considering vagueness in expert judgments.

This interpretation of fuzzy D + R and D – R values in *Table 13* provides a clearer picture of how each factor behaves in relation to others under fuzzy conditions.

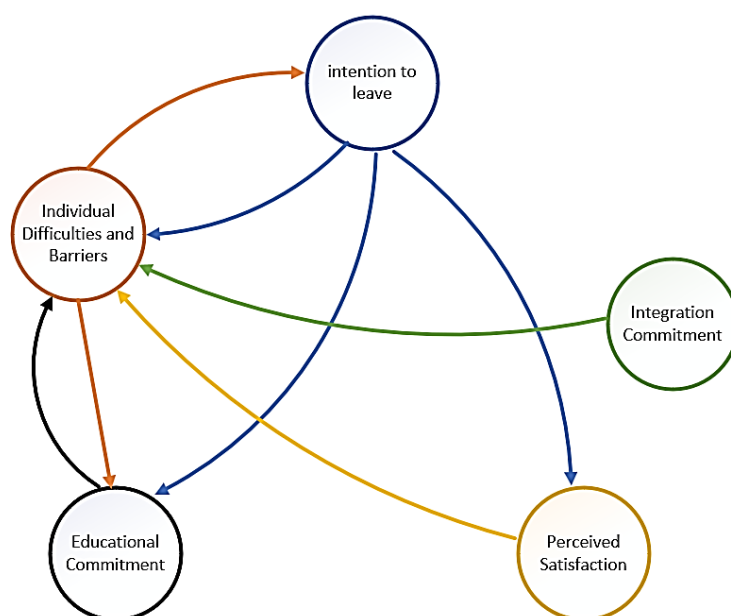


Fig. 2. Fuzzy network diagram of factors influencing international student retention using fuzzy DEMATEL method.

Fig. 2 visually maps the relationships between the factors, with the thickness and direction of fuzzy arrows illustrating both the strength and direction of influence. This diagram highlights the central role of intention to leave as a key influencer in the system while showing how individual difficulties/barriers depend on other factors. This analysis shows that intention to leave exerts a strong fuzzy influence on the system, while individual difficulties and barriers are highly reactive under uncertain conditions, depending on other factors for their dynamics.

The results emphasize that intention to leave and individual difficulties and barriers are the most critical elements to address to improve international student retention, even when considering fuzzy evaluations. Integration commitment and educational commitment, while less prominent, still play significant roles and should not be overlooked.

By applying the fuzzy DEMATEL method, we capture the complexity and uncertainty in the interdependencies and causal relationships among key factors, providing a flexible yet structured framework for developing effective retention strategies under conditions of uncertainty.

4.6.1 | Policy implications

The findings from this study, supported by both traditional statistical methods and the fuzzy DEMATEL analysis, offer significant insights for policymakers and higher education institutions aiming to improve the retention of international students. Retaining international students is not only economically beneficial but also culturally, academically, and socially vital for universities and the broader community. Based on our analysis, several key policy implications emerge.

Addressing individual difficulties and barriers

The fuzzy DEMATEL analysis identified individual difficulties and barriers as a highly reactive factor ($D - R = -0.637$), meaning that it is heavily influenced by other factors, such as perceived satisfaction and educational commitment. This underscores the importance of creating targeted interventions to mitigate individual challenges, such as financial constraints, language barriers, and cultural adjustments.

Policymakers should consider enhancing support systems, including: 1) Financial aid programs: expanding scholarships or flexible payment options for tuition and living expenses to ease the financial burdens of

international students, and 2) Language and cultural support: offering intensive language programs and cross-cultural orientation sessions to help students overcome integration challenges.

These interventions will not only reduce individual difficulties but also positively impact retention by improving student satisfaction and educational commitment, as revealed by the strong correlations between these variables.

To operationalize these recommendations, universities can develop targeted financial aid programs that consider the unique needs of international students, such as flexible payment plans and emergency financial support. Language and cultural support initiatives can include mandatory pre-semester language workshops and ongoing cultural adaptation sessions facilitated by experienced staff and student mentors. These programs will foster a sense of belonging and support, thereby enhancing students' ability to overcome individual challenges.

Enhancing educational and social integration

According to both the regression and fuzzy DEMATEL results, educational commitment ($\beta = -0.262$) and integration commitment ($\beta = -0.005$) have a direct effect on students' intentions to leave, with educational commitment showing a significant negative impact. This suggests that institutions should prioritize strategies that promote both academic engagement and social integration.

Policies should focus on: 1) Personalized academic support: implementing tailored academic advising and mentoring programs to increase students' educational commitment and reduce their intention to leave. Creating opportunities for international students to engage in research projects or career-oriented initiatives will also strengthen their academic ties to the institution, and 2) Social integration initiatives: facilitating student organizations, cultural exchange programs, and peer-mentoring systems to help international students build social connections and a sense of belonging within the university community.

These measures align with the fuzzy DEMATEL analysis, where integration commitment and educational commitment were found to have strong positive relationships with perceived satisfaction. Stronger integration and academic commitment directly contribute to higher satisfaction levels, ultimately reducing the intention to leave.

Universities can implement these strategies by establishing personalized academic advising programs that connect students with faculty mentors who understand the unique challenges faced by international students. Additionally, social integration initiatives such as cultural exchange events, international student clubs, and peer mentorship programs can be developed to create a welcoming community. These measures should be designed with feedback from international students to ensure relevance and effectiveness.

Improving perceived satisfaction

The regression results reveal that perceived satisfaction ($\beta = -0.409$) has the strongest negative impact on students' intention to leave, highlighting it as a critical factor for retention. The fuzzy DEMATEL analysis further supports this, showing that perceived satisfaction is both influenced by and has a strong influence on educational and integration commitment ($D + R = 1.486$).

Therefore, universities should: 1) Optimize student services: enhance services such as academic counseling, health and wellness support, housing, and administrative assistance. These services directly improve students' perceived satisfaction with their university experience, and 2) Continuous feedback mechanisms: implement regular surveys and feedback channels to gauge student satisfaction and promptly address any issues related to academics, social life, or institutional services.

Improving overall satisfaction will have a cascading positive effect on other critical retention factors, such as reducing individual difficulties and increasing commitment levels.

To enhance perceived satisfaction, universities can optimize their student support services by expanding academic counseling hours, offering accessible health and wellness resources, and streamlining administrative

processes to minimize bureaucratic hurdles. Continuous feedback mechanisms, such as regular student surveys and open feedback sessions, should be institutionalized to identify and address any areas of dissatisfaction promptly.

Targeting demographic and institutional challenges

The fuzzy DEMATEL analysis identified key demographic factors such as family income, distance from the university, and visa or residence permit issues as critical influences on student retention.

Policy strategies should include: 1) Tailored support for vulnerable groups: provide additional financial and logistical support to students from lower-income families or those facing residence and visa challenges. This may include offering housing options closer to campus, streamlining visa processes, or creating partnerships with local governments for residency assistance, and 2) Focused retention strategies for at-risk students: develop early intervention programs that identify students at risk of leaving based on demographic factors. Machine learning models, such as the Random Forest algorithm used in this study, can help predict students' likelihood of leaving based on these factors and guide targeted retention efforts.

Institutions can address these demographic challenges by creating tailored support systems for vulnerable student groups. For example, additional financial grants or subsidies can be provided to students from lower-income backgrounds, and partnerships with local authorities can help streamline residence and visa-related processes. Early identification programs using predictive analytics can help institutions preemptively identify students who may be at risk of dropping out and offer tailored support.

Strategic importance of intention to leave

The fuzzy DEMATEL analysis positioned intention to leave as the most influential factor ($D - R = 0.637$), indicating that it is a primary driver of the retention system. Therefore, institutions must develop proactive strategies to prevent international students from considering departure in the first place.

This could involve: 1) Early warning systems: using data analytics to identify early signs of disengagement, such as declining academic performance or reduced participation in campus activities and providing timely support, and 2) Career development opportunities: offering internships, co-op placements, and career counseling to ensure that international students see clear pathways to success post-graduation, which can reduce their intention to leave.

To effectively mitigate the intention to leave, universities should invest in early warning systems powered by data analytics that monitor student engagement metrics such as class attendance, assignment submissions, and participation in campus activities. Proactive measures such as outreach by academic advisors and student support teams should be deployed when signs of disengagement are detected. Additionally, career development programs that connect international students with internship and co-op opportunities can reinforce their commitment to remain enrolled by demonstrating clear post-graduation pathways.

5 | Conclusion

This study offers a comprehensive investigation into the factors influencing international student retention in Turkish higher education institutions. By employing both traditional statistical methods and machine learning algorithms, several critical variables that affect students' decisions to either stay or leave were identified. The findings indicate that individual difficulties and barriers, such as financial and personal challenges, significantly increase the likelihood of students intending to leave. Conversely, factors such as integration commitment, educational commitment, and perceived satisfaction play a crucial role in reducing this intention.

Demographic variables, including family income, residential distance, and visa or residence permit issues, also emerged as key predictors of retention. These insights emphasize the importance of addressing not only academic and social integration but also personal and financial challenges faced by international students.

The machine learning analysis, particularly the Random Forest algorithm, was the most effective in predicting retention outcomes, highlighting the potential of AI tools in shaping university retention strategies. By ranking the most important features, the study identified happiness at the university, academic commitment, and perceived satisfaction as the top factors influencing student retention. The analysis highlighted that the Random Forest algorithm achieved the highest accuracy, attributed to its ensemble approach that effectively handles complex and non-linear relationships within the data. This makes it a robust choice for studies involving multiple interdependent factors, such as student retention.

Through the application of the fuzzy DEMATEL method, the study identified intention to leave and individual difficulties and barriers as the most significant factors impacting international student retention based on their high ($D + R$) values. The net influence analysis ($D - R$) further revealed that intention to leave serves as a key driver within the system, while individual difficulties and barriers are more dependent, reacting strongly to other factors. These findings align with the hypothesis testing results, reinforcing the understanding of how key factors like intention to leave, individual challenges, and levels of commitment interrelate.

In conclusion, universities and policymakers must adopt a holistic approach to improving student retention, which includes providing financial support, enhancing academic and social integration, and addressing the personal challenges faced by international students. Institutions that implement these strategies effectively will not only improve retention rates but also cultivate a more inclusive and supportive academic environment.

Future research could build on these findings by exploring additional demographic variables or applying these models in different cultural and institutional contexts to verify their generalizability. Longitudinal studies could be conducted to observe retention patterns over time and analyze how changes in institutional policies impact student retention. Researchers could also integrate qualitative methods, such as interviews and focus groups, to gain deeper insights into the personal experiences of international students. Expanding the analysis to include other psychological variables, such as motivation and resilience, would further enrich the understanding of factors influencing retention. Additionally, applying advanced machine learning techniques to larger datasets could refine predictive models and improve the accuracy of retention forecasts. Ultimately, improving international student retention benefits not only the students themselves but also contributes to the economic, cultural, and academic advancement of the host country.

Conflict of Interest

The author declares no conflicts of interest related to this research. There has been no financial, personal, or professional involvement that could influence the work reported in this paper.

Data Availability

The datasets generated and analyzed during the current study are not publicly available due to institutional restrictions but are available from the corresponding author on reasonable request and with permission from Istinye University.

Author Contributions

The sole author contributed to all aspects of the study, including the conceptualization, methodology, data collection, analysis, and manuscript preparation.

References

- [1] Agasisti, T., & Bertoletti, A. (2022). Higher education and economic growth: A longitudinal study of European regions 2000–2017. *Socio-economic planning sciences*, 81, 100940. <https://doi.org/10.1016/j.seps.2020.100940>
- [2] Bertoletti, A., Berbegal-Mirabent, J., & Agasisti, T. (2022). Higher education systems and regional economic development in Europe: A combined approach using econometric and machine learning methods. *Socio-economic planning sciences*, 82, 101231. <https://doi.org/10.1016/j.seps.2022.101231>
- [3] Bülbül, T. (2012). Yükseköğretimde okul terki: Nedenler ve çözümler. *Eğitim ve bilim*, 37(166). <http://eb.ted.org.tr/index.php/EB/article/view/1490>
- [4] Farrugia, C., & Bhandari, R. (2020). Global trends in student mobility. In *The international encyclopedia of higher education systems and institutions* (pp. 560–568). Springer. https://doi.org/10.1007/978-94-017-8905-9_231%0A%0A
- [5] Contini, D., & Salza, G. (2020). Too few university graduates, inclusiveness and effectiveness of the Italian higher education system. *Socio-economic planning sciences*, 71, 100803. <https://doi.org/10.1016/j.seps.2020.100803>
- [6] Bao, T., Liu, Y., Yang, Z., Wu, S., & Yan, Z. (2024). Evaluating sustainable service quality in higher education from a multi-stakeholder perspective: An integrated fuzzy group decision-making method. *Socio-economic planning sciences*, 92, 101849. <https://doi.org/10.1016/j.seps.2024.101849>
- [7] OCED. (2016). *Education at a glance 2016 OECD indicators*. https://www.oecd.org/en/publications/2016/09/education-at-a-glance-2016_g1g6a8d2.html
- [8] Gök, S., & Akar, H. (2023). Being a pioneering alternative school in turkey: values, pedagogical practices, and challenges. *Eğitim ve bilim*, 48(216). <http://doi.org/10.15390/eb.2023.11778>
- [9] Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition*. University of Chicago Press.
- [10] Gülşen, F. U., Aydın, A., & Gizir, S. (2022). Dropout from higher education in Turkey: A qualitative study. *Yükseköğretim Dergisi*, 12(1), 39–50. <https://doi.org/10.2399/yod.21.587615>
- [11] Tinto, V. (1997). Classrooms as communities: Exploring the educational character of student persistence. *The journal of higher education*, 68(6), 599–623. <https://doi.org/10.1080/00221546.1997.11779003>
- [12] Pascarella, E. T., & Terenzini, P. T. (1980). Predicting freshman persistence and voluntary dropout decisions from a theoretical model. *The journal of higher education*, 51(1), 60–75. <https://doi.org/10.1080/00221546.1980.11780030>
- [13] Bean, J. P., & Metzner, B. S. (1985). A conceptual model of nontraditional undergraduate student attrition. *Review of educational research*, 55(4), 485–540. <https://doi.org/10.3102/00346543055004485>
- [14] Astin, A. W. (1999). Student involvement: A developmental theory for higher education. *Journal of college student development*, 40(5), 518–529. <https://psycnet.apa.org/record/1999-01418-006>
- [15] National Center for Education Statistics (NCES). (2019). *Role in monitoring and supporting student success through data collection and analysis*. <https://nces.ed.gov/>
- [16] Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3), 338–353. [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X)
- [17] Zimmermann, H.-J. (2011). *Fuzzy set theory—and its applications*. Springer Science & Business Media.
- [18] Uluçay, V., & Şahin, M. (2024). Intuitionistic fuzzy soft expert graphs with application. *Uncertainty discourse and applications*, 1(1), 1–10. <https://uda.reapress.com/journal/article/view/16>
- [19] Priyadarshini, S., Nayak, S., & Panigrahi, P. K. (2022). Fuzzy finite element solution of uncertain convection-diffusion heat transfer for a rectangular plate. <https://doi.org/10.48550/arXiv.2205.10132>
- [20] George J, K., & Bo, Y. (2008). Fuzzy sets and fuzzy logic, theory and applications. <https://digilib.uin-suka.ac.id/id/eprint/7049/>
- [21] Chang, B., Chang, C. W., & Wu, C. H. (2011). Fuzzy DEMATEL method for developing supplier selection criteria. *Expert systems with applications*, 38(3), 1850–1858. <https://doi.org/10.1016/j.eswa.2010.07.114>

- [22] Lin, K. P., Tseng, M. L., & Pai, P.-F. (2018). Sustainable supply chain management using approximate fuzzy DEMATEL method. *Resources, conservation and recycling*, 128, 134–142. <https://doi.org/10.1016/j.resconrec.2016.11.017>
- [23] Gül, S. (2022). Picture fuzzy extension of DEMATEL and its usage in educational quality evaluation. In *New perspectives in operations research and management science: essays in honor of fusun ulengin* (pp. 471–497). Springer. https://doi.org/10.1007/978-3-030-91851-4_18
- [24] Torkzadeh, J., Pamucar, D., & Niroomand, S. (2023). Interval-valued pythagorean fuzzy extension of dematel for prioritizing and casualty analysis of environmental criteria of organizational behaviour in higher education sector. *Informatica*, 1–26. <https://doi.org/10.15388/23-INFOR541>
- [25] Ansari, Z. N., Kant, R., & Shankar, R. (2020). Evaluation and ranking of solutions to mitigate sustainable remanufacturing supply chain risks: a hybrid fuzzy SWARA-fuzzy COPRAS framework approach. *International journal of sustainable engineering*, 13(6), 473–494. <https://doi.org/10.1080/19397038.2020.1758973>
- [26] Ayouni, S., Jamel, L., Hajjej, F., & Maddeh, M. (2019). *A hybrid fuzzy dematel-ahp/vikor method for LMS selection* [presentation]. Processing of the ecel 2019 18th european conference on e-learning (pp. 47–55). [https://books.google.com/books?hl=en&lr=lang_en&id=ZMjBDwAAQBAJ&oi=fnd&pg=PA47&dq=%5B24%5D+%09Ayouni,+S.,+Jamel,+L.,+Hajjej,+F.,+%26+Maddeh,+M.+\(2019\).+A+hybrid+fuzzy+dematel-ahp/vikor+method+for+lms+selection+%5Bpresentation%5D.+Processing+of+the+ecel+2019+18th+europe an+conference+on+e-learning+\(pp.+47%E2%80%9355\).&ots=NOK2pXZdHM&sig=Ud7RKRgnTCV7KkluPxSwnn1XdAY](https://books.google.com/books?hl=en&lr=lang_en&id=ZMjBDwAAQBAJ&oi=fnd&pg=PA47&dq=%5B24%5D+%09Ayouni,+S.,+Jamel,+L.,+Hajjej,+F.,+%26+Maddeh,+M.+(2019).+A+hybrid+fuzzy+dematel-ahp/vikor+method+for+lms+selection+%5Bpresentation%5D.+Processing+of+the+ecel+2019+18th+europe an+conference+on+e-learning+(pp.+47%E2%80%9355).&ots=NOK2pXZdHM&sig=Ud7RKRgnTCV7KkluPxSwnn1XdAY)
- [27] Dinçer, H., & Yüksel, S. (2020). Defining the strategic impact-relation map for the innovative investments based on IT2 fuzzy DEMATEL: Evidence from the European tourism sector. In *Handbook of research on creating sustainable value in the global economy* (pp. 103–124). IGI Global. <https://doi.org/10.4018/978-1-7998-1196-1.ch007>
- [28] Jeong, J. S., & González-Gómez, D. (2020). Adapting to PSTs' pedagogical changes in sustainable mathematics education through flipped E-Learning: Ranking its criteria with MCDA/F-DEMATEL. *Mathematics*, 8(5), 858. <https://doi.org/10.3390/math8050858>
- [29] Thavi, R. R., Narwane, V. S., Jhaveri, R. H., & Raut, R. D. (2022). To determine the critical factors for the adoption of cloud computing in the educational sector in developing countries – a fuzzy DEMATEL approach. *Kybernetes*, 51(11), 3340–3365. <https://doi.org/10.1108/K-12-2020-0864>
- [30] Jeong, J. S., González-Gómez, D., & Yllana-Prieto, F. (2022). Multi-criteria decision analysis and fuzzy-decision-making trial and evaluation laboratory (MCDA and F-DEMATEL) method for flipped and sustainable mathematics teaching as a real-life application. In *Real life applications of multiple criteria decision making techniques in fuzzy domain* (pp. 105–123). Springer. https://doi.org/10.1007/978-981-19-4929-6_5
- [31] Banihashemi, S. A., Khalilzadeh, M., & Pamucar, D. (2023). Analysis of the organizational knowledge management system of nurse education centers with hybrid Fuzzy DEMATEL-network DEA method. *International journal of knowledge-based and intelligent engineering systems*, 27(4), 395–406. <https://doi.org/10.3233/KES-230458>
- [32] Revalde, H., Almerino Jr, P., Rodriguez, M., Hortelano, P., Revalde, M., Chan, C., ... others. (2024). A q-Rung Orthopair Fuzzy Set Extension of the DEMATEL and Its Application in the Education Sector. *Advances in fuzzy systems*, 2024(1), 5593966. <https://doi.org/10.1155/2024/5593966>
- [33] Amro, D. K., & Dawoud, H. (2024). Influencing factors of spatial ability for architecture and interior design students: a fuzzy DEMATEL and interpretive structural model. *Buildings*, 14(9), 2934. <https://doi.org/10.3390/buildings14092934>
- [34] Basaran, M. A. (2012). Calculating fuzzy inverse matrix using fuzzy linear equation system. *Applied soft computing*, 12(6), 1810–1813. <https://doi.org/10.1016/j.asoc.2012.01.005>
- [35] Gürbüz, S., & Şahin, F. (2014). Sosyal bilimlerde araştırma yöntemleri. *Ankara: Seçkin Yayıncılık*, 271.